Elsevier Editorial System(tm) for Journal of Chromatography B

Manuscript Draft

Manuscript Number: JCB-16-108R1

Title: Evaluation of the Technical Variations and the Suitability of a Hydrophilic Interaction Liquid Chromatography-High Resolution Mass Spectrometry (ZIC-pHILIC-Exactive Orbitrap) for Clinical Urinary Metabolomics Study

Article Type: Full Length Article

Keywords: clinical metabolomics; Hydrophilic Interaction Liquid Chromatography-High Resolution Mass Spectrometry (HILIC-HRMS); untargeted metabolite profiling; technical variation; isotope labelled amino acids

Corresponding Author: Dr. Tong Zhang,

Corresponding Author's Institution: Glasgow University

First Author: Tong Zhang

Order of Authors: Tong Zhang; David G Watson

Dear Editor(s):

We would like to submit the manuscript titled “Evaluation of the Technical Variations and the Suitability of a Hydrophilic Interaction Liquid Chromatography-High Resolution Mass Spectrometry (ZIC-pHILIC-Exactive Orbitrap) for Clinical Urinary Metabolomics Study” to Journal of Chromatography B.

Currently LC-MS based untargeted metabolite profiling is the most widely used strategy for biomarker discovery in clinical metabolomics. However, it is the fact that the biomarkers reported in one research study are often invalid in another using different analytical method. Poor method reliability and robustness is considered as one of the reasons, but only a few investigation studies have been reported in respect of this concern and most of them used reversed phase chromatography.

As following up our previous studies on development and optimisation of Hydrophilic Interaction Liquid Chromatography-High Resolution Mass Spectrometry (HILIC-HRMS) for untargeted metabolite profiling, here using a series of human urine samples, we investigated the technical variations of this method and its suitability for supervised multivariate analysis

By comparison of LC-MS data generated by the samples (injections) prepared in different ways, we observed that the technical variation introduced by sample preparation was more significant than the current HILIC-HRMS method. Therefore, it is necessary to prepare and measure multiple QC samples with the real samples for assessment of data deviations introduced in sample preparation. By using a series of human urine samples spiked with isotope labelled amino acids at varied concentrations, to the best of our knowledge, at the first time, we demonstrated the suitability of a LC-MS method for untargeted metabolite profiling in supervised multivariate analysis ((e.g. Orthogonal Partial Linear Square: OPLS). The robustness of the current HILIC method was evaluated by manipulation of several key LC parameters that may affect the reproducibility in practical use for long term/batch-to-batch analysis. Some influential parameters were determined by Principal Component Analysis (PCA) of the retention times obtained under varied LC conditions. One of our important findings is the high reproducibility among different columns in chromatographic

performance of this polymer based zwtterionic stationary phase in HILIC mode, which allows column replacement for high sample throughput analysis among multiple batches.

Based on the observation in this study, this HILIC-HRMS (ZIC-pHILIC coupled with Exactive Orbitrap) method has been demonstrated to be reliable and suitable for biomarker discovery in clinical metabolomics.

Regards

Tong Zhang

22/01/16

Dear Editor(s):
Please find our replies to the reviewers' comments below in red.

Regards

Tong Zhang

Reviewer #1: Experimental methods were not accessible for review from author's prior publications. I suggest that the LC and MS experimental conditions be listed for convenience of the reader.
For convenience the LC-MS conditions have been added to the current paper.

Reviewer #3: In this paper Zhang and Watson expand upon the excellent work performed on Zic-pHILIC chromatography published by validating the robustness of the assay when analyzing urine. This work is well referenced and well written. Several issues of note:
1. I had a difficult time reading the figures due to poor graphics resolution in the pdf.
All the original figures are made with high resolution. The poor resolution was caused by the conversion to the PDF format when the figures were uploaded to the JCB website.

2. I would like to see a spread sheet of the analysis order to fully evaluate the robustness testing, as it is now I am not absolutely sure of how to evaluate the data.
Good point, as suggested Table 1 has been remade and the relevant description has also been added into section 2.3.

3. I would like to see a 5 by 5 analysis performed using the final assay conditions. That is run 25 samples total, 5 samples a day, 5 days in a row to determine robustness testing.
We don't think the assessment of intermediate precision in this way is applicable to metabolomics studies which are run continuously over perhaps 5 days rather than batch-wise and are thus open ended. Thus any degradation in performance comes from cumulative effects such as build-up of matrix components in the background or the LC-MS source becoming dirty. The method for controlling fall off in performance during batch runs is the use of a pooled sample which is run at intervals throughout the run. The current study was aimed at assessing the robustness of the method in withstanding variations in factors such as column temperature, mobile phase pH and the variation in column batch which occurs between batches which might be run weeks or months apart. The good intermediate precision of this HILIC-HRMS method has been demonstrated in our previous publications using the data from pooled QC samples run with the real samples. Thus the purpose of the current paper is to recommend the method at the level of reproducibility i.e. it could be adopted by another laboratory without loss of performance.

Evaluation of the Technical Variations and the Suitability of a Hydrophilic Interaction Liquid Chromatography-High Resolution Mass Spectrometry (ZIC-pHILIC-Exactive Orbitrap) for Clinical Urinary Metabolomics Study

Tong Zhang* and David G Watson

Strathclyde Institute of Pharmacy and Biomedical Sciences, 161, Cathedral Street, Glasgow, Scotland, UK, G4 0RE.

*Corresponding author:

Tong Zhang,

Wolfson Wohl Cancer Research Centre, Institute of Cancer Sciences, University of Glasgow

Switchback Road

Glasgow, UK

G61 1QH,

E-mail: tong.zhang@glasgow.ac.uk

Abstract

Although many hydrophilic interaction liquid chromatography-high resolution mass spectrometry (HILIC-HRMS) methods have been developed and applied for untargeted metabolite profiling in clinical metabolomics, according to the literature, the suitability of these HILIC-HRMS methods has not been fully evaluated with respect to their performance when they are subjected to statistical analysis. In this study, using a series of human urine samples we investigated the effect of technical variations on multivariate and univariate analysis of the data collected using a previously developed HILIC-HRMS method for untargeted urinary metabolite profiling in clinical metabolomics. The technical variation introduced by sample preparation was more significant than that produced by the HILIC-HRMS method. By using an orthogonal partial least squares (OPLS) model, subtle fold-changes were accurately measured in the urine samples spiked with ^{13}C and ^{15}N isotope labelled amino acids at different concentrations. The robustness of this HILIC method was also evaluated by analysing the obtained data from a single urine sample following manipulation of several primary LC parameters. High reproducibility in the chromatographic performance of three ZIC-pHILIC columns with different batch numbers indicated the reliability of the polymer based zwitterionic stationary phase allowing column replacement without compromising the performance of the method.

Keywords: clinical metabolomics; Hydrophilic Interaction Liquid Chromatography-High Resolution Mass Spectrometry (HILIC-HRMS); untargeted metabolite profiling; technical variation; isotope labelled amino acids

1 Introduction

In clinical metabolomics, untargeted metabolite profiling is designed for identification and comparative analysis of hundreds or thousands of metabolites in clinical samples e.g. urine, blood and tissues. [1-6] Associated with multivariate analysis (MVA), it offers the opportunity to compare/differentiate the metabolic fingerprints of clinical samples and to discover individual biomarker candidates for identifying particular pathological states. By following this strategy, untargeted metabolite profiling of biofluid samples has been used for the development of new techniques potentially applicable in clinical use, such as disease diagnosis and prognosis, therapeutic treatment evaluation and disease screening. [1-6] However, up until now no metabolic biomarkers have been reported for practical use in clinics, with the exception perhaps of a method for the typing of the rare conditions xanthuria types I and II,[7, 8] especially for common diseases such as cancer. This results from the fact that poor reproducibility has been reported for biomarkers between research groups where often insufficient cohorts have been involved in these studies.[9, 10] For instance, sarcosine was reported as a biomarker metabolite correlated with prostate cancer progression in 2009, [11] but this was followed by many contradictory results found by different research groups using different analytical methods.[12] Besides vague experimental designs and inappropriate data analysis/interpretation, an unreliable and/or unsuitable analytical method can be considered as one of the principal reasons for poor biomarker reproducibility.[13, 14]

Nowadays, Mass Spectrometry (MS) coupled with Liquid Chromatography (LC) is the most widely employed analytical platform for untargeted metabolite profiling of biological samples.[2, 3, 6, 15] Many LC-MS method development and comparison studies have been

reported.[16-21] In order to approach the maximum metabolome coverage, the number of detected metabolites is of the greatest concern in these studies. The quality of the chromatographic signals should be addressed because poor/irregular chromatographic peaks for metabolites could lead to false outcomes in data processing which can be easily overlooked by bioinformaticians and biologists who may be unaware of the impact of chromatography on data quality.[17] The other critical point, especially for clinical metabolomics, is the LC-MS method stability for long-term experiments (intra- and inter-batch) and its robustness in withstanding to small changes in LC-MS conditions which are inevitable for batch-to-batch analysis of samples. Such changes can result from factors such as LC column replacement, variations in mobile phase composition, the accuracy of the gradient formation by the HPLC and MS system maintenance. Although some data processing algorithms can be applied for reducing the deviations introduced in data acquisition, the quality of the raw data should be addressed as a priority, thus the technical variations caused by sample preparation and data acquisition in an analytical method should be evaluated for determination of its reliability in clinical metabolomics studies.

Recently a few studies have been reported with respect to the stability of LC-MS platforms for untargeted metabolite profiling using unsupervised MVA (e.g. Principal Component Analysis: PCA) of Quality Control (QC) samples/injections.[18, 22-25] In real clinical metabolomics studies, however, supervised MVA (e.g. Orthogonal Partial Linear Squares-Discriminant Analysis: OPLS-DA) is much more effective for uncovering biomarkers due to the large matrix effect of clinical samples which is more likely orthogonal to the genuine pathological changes in such samples. According to our literature review, no studies have demonstrated the suitability of an LC-MS method in such circumstances by using supervised

MVA. In addition, there is increasing evidence that hydrophilic interaction liquid chromatography (HILIC) is more advantageous than, or at least highly complementary to, reversed phase (RP) LC in untargeted polar metabolite profiling coupled with HRMS.[9, 26]. In our previous studies, we have systematically demonstrated the superiority of a HILIC-HRMS (ZIC-pHILIC-Exactive Orbitrap) method in metabolite coverage and chromatographic performance for untargeted profiling of polar metabolites in biological samples.[16, 17] In this study, in order to further evaluate the reliability and suitability of this method for clinical metabolomics, evaluation tests were performed using a series of spiked samples prepared from one single urine sample and the generated LC-MS data were assessed by both unsupervised and supervised MVA (PCA and OPLS) as well as reproducibility in peak area and retention time for individual LC-MS signals. Several key LC parameters, which might potentially introduce systematic variation to the measurements between batches and/or laboratories, were also evaluated in order to test the robustness of the HILIC method.

2 Material and methods

2.1 Chemicals

HPLC grade acetonitrile (ACN) was obtained from Fisher Scientific, UK. Ammonium carbonate, formic acid and ammonium hydroxide solution (28-30%) were purchased from Sigma-Aldrich, UK. HPLC grade water was produced by a Direct-Q 3 Ultrapure Water System from Millipore, UK. The standard mix of sixteen ^{13}C and ^{15}N isotope fully labelled amino acids (Electronic Supplementary Information.1: ESI.1) was purchased from Sigma-Aldrich, UK (487910-500mg) and a stock solution was prepared at the concentration of 1.5mg/ml in the mixture of ACN and water (1:1). Because this standard mixture was extracted from blue-green algae, it is believed that it contains some other metabolites which are not heavy

isotope labelled. The stock solution was diluted 10 times with ACN only to prepare the solution (150µg/ml) named as dilution factor (DF) 1. This DF1 solution was further diluted 5, 10, 15 and 20 times with ACN to obtain the solutions at the concentration of 30, 15, 10 and 7.5µg/ml and correspondingly named as DF5, DF10, DF15 and DF20, respectively which were used as the ACN solutions required in the following urine sample preparation.

2.2 Urine sample preparation

One hundred millilitres of human urine was obtained with ethics permission from our previous study.[27] A series of samples were prepared from this single urine sample as schematically shown in Figure.1. The replicate samples were named in the following figures with the urine volumes used in the preparation or the dilution factors of isotope labelled amino acids in ACN (see Figure.1). After centrifugation at 10,000 RPM for 5 minutes the clear solutions were transferred into LC auto-sampler vials and randomly sequenced for the HILIC-HRMS experiment with QC injections (named as QCI) from a single 200-QC sample. Including blank injections this whole experiment lasted approximately 56 hours. For HILIC method robustness tests one single sample of QC-200 was used.

2.3 HILIC-HRMS experiment

The HILIC-HRMS experiments were carried out on a Dionex Ultimate 3000 HPLC system coupled with an Exactive (Orbitrap) mass spectrometer (Thermo Fisher Scientific, Germany). Three ZIC-pHILIC columns (150×4.6mm, 5µm) used in this study were purchased from HiChrom (Reading UK). For evaluation of the technical variations in this analytical method, all the HILIC-HRMS parameters were set to the values optimised in our previous studies [16, 17] as described below: the mass scanning range was m/z 75–1200, the needle voltage was

4.5 kV in positive mode and -4.0 kV in negative mode, the capillary temperature was 320 °C and the sheath and auxiliary gas flow rates were 50 and 17 arbitrary units, respectively. The mobile phase used was 20 mM ammonium carbonate buffer (pH 9.2) (A) and pure ACN (B); the flow rate was 300 µl min^{-1}. The gradient was programed as follows: 0 min 20% A 80% B to time 30 min 80% A 20% B 31 min 8 % B 36 min 8 % B 37 min 80% B 46 min 80% B. The injection volume was 10 µl and the sample tray temperature was controlled at 12 °C during the measurement.

For HILIC method robustness tests each primary LC parameter was examined at three levels and the manipulation is detailed in Table 1. Only one single parameter was changed under each condition with the others fixed at the standard settings and one blank injection was made for equilibration purpose followed by 7 consecutive injections of a single urine sample (QC-200 in Figure 1). The manipulations of the column temperature and the sample tray temperature were carried out by programmed LC methods on a new ZIC-pHILIC column in a sequence of (CT10, Std, CT40, ST5 and ST20) and the mobile phases and the columns were manually changed according to the protocol (Table 1). It should be aware that some parameter setting manipulations in this study were beyond the definition of routine robustness test, the main purpose was to stimulate the possible condition deviations in inter-batch analysis and evaluate the response of this HILIC-HRMS method to the changes.

2.4 Data processing

The Xcalibur raw data was converted into .mzXML files using ProteoWizard and imported into the MZMine 2.10 [28] for peak extraction and alignment. The detailed procedure and the parameter settings were the same as in our previous study [29] except for retention time tolerance window in alignment. Considering the possibility of significant changes in

retention time for some metabolites, the tolerance window was enlarged to 20% for alignment of the samples measured under varying LC conditions. The output .CSV file was imported into an in-house macro (Microsoft Excel 2010) for subtraction of the signals from blank injections. Putative annotation was also conducted in the macro by searching the accurate mass of M±H with a tolerance window of 3 ppm in our in-house database.[16] In biomarker discovery studies the US Food and Drug Administration (FDA) guidance allows up to 30% relative standard deviation (RSD) for signal variation among QC samples and Dunn et al suggested 20% for RPLC-HRMS. [30] In this study considering relatively poor chromatographic stability of HILIC compared to RPLC, we chose 25% RSD as the acceptance tolerance for signal variation. In order to assess the unrepeatability caused by sample preparation and data acquisition in this method, no RSD filter was applied on signal variation in data analysis for reliability evaluation.

2.5 Statistical analysis

SIMCA version 14.0 (Umetrics, Sweden) was used for MVA including PCA and OPLS. In order to prevent the model being dominated by several highly abundant urinary metabolites (e.g. creatinine and hippurate in ESI positive and negative mode respectively), unit variance (UV) scaling was used for all PCAs, but for OPLS Pareto scaling was used to facilitate the discovery of the spiked “biomarkers” (isotope labelled AAs) in the loading plot. No data normalisation and transformations were performed because all the samples were prepared from one single urine sample. Statistical functions in Microsoft Excel 2010 were used for basic calculations e.g. RSD and R^2 value in linear regression analysis.

3 Results and discussion

3.1 Evaluation of the technical variations in sample preparation and data acquisition

A reliable LC-MS method for untargeted metabolite profiling is expected to determine biological variations among samples as opposed to technical variations which are inevitably introduced by sample preparation and data acquisition. Although this study might seem rather contrived it is designed to illustrate in a simple fashion the steps that occur in a typical LC-MS based metabolomics study and illustrate how variations might arise which could confuse data interpretation. As shown in Figure 1, in order to test the technical variations in the current analytical method, replicate urine samples were prepared with varied total volumes or dilution factors and one of the QC-200 samples was injected 17 times throughout the whole run sequence. The LC-MS data generated was processed for the statistical analysis which followed. After removing the background signals from the blanks, 2754 and 3017 LC-MS features were detected in the samples in the ESI positive and negative modes respectively. PCA was performed using the combined dataset without any filtration according to signal repeatability among QC injections (QCI). It can be observed in Figure 2A that the samples are primarily classified along the first principal component (PC1) representing dilution factor and the first four QCIs are clearly apart from others with the same urine dilution factor 5. After excluding the first four QCIs, in Figure 2B the clustering of each group was improved no marked separation of the samples with a dilution factor of 5 can be observed. Compared to RPLC, HILIC is considered to require more time for system equilibration in gradient elution which is usually achieved by consecutively injecting a single QC sample prior to measurement of real samples. [25, 30] In the current case, four consecutive injections of a QC sample appear necessary for sufficient equilibration of this HILIC-HRMS system. In order to further distinguish the variations caused by sample

preparation and instrument instability, as shown in Figure 2C the PCA score plot was generated by only using the samples prepared with dilution factor of 5. Low R2X (0.375) and Q2 (0.307) scores were obtained by a combination of PC1 and PC2 indicating a poor fit for the model. This can be explained by the poor correlations in this dataset which can be attributed to the high similarity of LC-MS data for these samples. Regardless of the poor fit of the model, it still can be seen in Figure 2C that the QCIs distribute along PC1 in a clear run time dependent manner and they are generally separated from most of other samples by the PC2. This observation implies that the LC-MS signal drift appears to be producing more serious technical variation than the sample preparation for the method. Interestingly, however, the opposite observation is reflected in Figure 2D. The samples prepared in the same way (QC-200: urine dilution factor=5 and total volume=1ml) generated more unrepeatable signals (RSD>25%) than multiple injections of a single sample (QCIs) even including the first four injections. Moreover, this number increased when a variable of urine volume, was used in the sample preparation but no clear classification can be observed in Figure 2C according to this variable. This phenomenon might be explained by the fact that MVA is more applicable for detection of systematic correlation/variation (dilution factors/signal drifting with time in the current case) rather than random deviation (sample preparation). Actually many LC-MS signals with poor RSDs among the QC-200 samples showed low but reproducible responses among QCIs (highlighted in green for the RSD% values in ESI.2) indicating that they are produced by low-abundance components in urine and can be reproducibly detected by this HILIC-HRMS platform and thus the variations are introduced by the sample preparation. Therefore, as addressed by Dunn et al [31], multiple individual QC samples (≥3) should be prepared and measured along with real samples not

only for monitoring the system stability of the instruments but also for assessing deviations for individual metabolites introduced in the sample preparation.

The samples prepared with varied total volumes or dilution factors showed a slight difference in the number of LC-MS signals with RSD below 25% except that a smaller dilution factor does improve the reproducibility in the measurement of low-abundance components in urine (see ESI.3A). It should be noticed that many of these were putatively annotated as oligopeptides (2-5 amino acids) based on accurate mass (highlighted in yellow for the RSD% values in ESI.2). The major concern for samples prepared with a small dilution factor is saturation of the mass spectrometry detection of high-abundance components in urine. Creatinine and hippurate were detected by the current method as the largest chromatographic peaks in ESI positive and negative mode respectively. By comparing the peak areas, significant differences could be observed according to the dilution factors but the correlation coefficient was only valid for hippurate and not for creatinine (ESI.3B) which indicates the reliability of this method for comparative analysis but a limitation in absolute quantification. This observation suggests that less dilution of urine may facilitate LC-MS detection of the low-abundance urinary metabolites while reducing the reliability of the absolute quantification for the high-abundance metabolites.

3.2 Method suitability evaluation using OPLS model

The metabolic perturbation caused by local pathogenicity (especially for early detection of chronic diseases such as cancer) on metabolome of human beings could be subtle thus this change in the metabolic signatures of clinical samples usually cannot be visualised by using unsupervised MVA (e.g. PCA). Therefore, supervised MVA (e.g. OPLS-DA) is often employed in clinical metabolomics studies. Here, we demonstrated the suitability of the current HILIC-

HRMS method in combination with OPLS analysis. As shown in Figure.1, ACN solutions containing sixteen isotope labelled (^{13}C and ^{15}N) amino acids were spiked into the urine samples with varying concentrations and together with the non-spiked urine samples (QC-200) the LC-MS data from these samples was used to prepare the MVA plots are shown in Figure 3. The DF1 samples were found as outliers in the original PCA plot because of the high concentration of spiked isotope labelled amino acids and thus they were excluded in the following analysis. As expected and shown in Figure 3A, under the dominating effect of the urinary metabolome the PCA model was unable to distinguish the subtle differences in the dataset caused by the spiking of isotope labelled amino acids. In contrast as shown in Figure 3B, the OPLS model accurately classified the samples according to the concentrations of the isotope labelled amino acids, even with small fold-changes, along the X axis (the correlated principal component). Large variations along the Y axis (the first orthogonally correlated principal component) can be observed in each group. This outcome was validated in comparison with the PCA model generated using the samples spiked into water thus removing the masking effect of the urinary metabolome. As shown in Figure 3C, the samples were naturally classified by the dilution factors for the isotope labelled amino acids along PC1. Figure 3D shows the loading plot of the OPLS model. The LC-MS signals generated by isotope labelled amino acids were identified in both ESI positive and negative modes and represented as solid red marks in the figure. As can be seen all of them associate with the X axis in a positive direction indicating highly positive correlation with the Y-variable (dilution factor). Due to the fact that the labelled amino acid mixture was extracted from blue-green algae; it contains some metabolites which are not naturally present in human urine, and many other hollow green marks can be also seen on the X axis showing highly positive correlation with dilution factor. On the other hand, the metabolites present in both human

urine and the blue-green algal extract exhibit a noticeable deviation along the Y axis such as uric acid and citric acid (solid green marks in Figure 3D). Significant ion suppression for the heavy isotope labelled amino acids by the corresponding light isotope signals in urine can be observed when comparing their peak areas with the same dilution factors in water (ESI.4). Apart from that observation, the correlation of the signal of each isotope labelled amino acid to dilution factor was generally good (the mean R^2 value calculated in linear regression analysis for all the isotope labelled amino acids was 0.989 ± 0.018 and 0.988 ± 0.027 in ESI negative and positive modes, respectively). However, considering the complexity of the urine metabolome for real clinical samples, care should be taken with regard to ion suppression of low intensity signals during the selection of biomarker candidates.

3.3 Robustness evaluation for LC parameters

Reproducible and reliable LC performance of intra- and more importantly inter-batch analysis is the key for chromatographic alignment of data during processing. As described in Table 1, four primary LC parameters were manipulated to simulate possible inter-batch deviations in LC conditions. The responses of the HILIC-HRMS method were demonstrated by PCA of the retention times of the 200 most intense LC-MS signals putatively annotated as real metabolites in ESI positive and negative ion mode. The retention times were manually defined based on the metabolite annotation in a few cases when alignment failed to find the peaks. Based on the distances to the cluster of usual standard conditions (Std) locating in the middle of the plot in Figure 4A, the order of effect on retention time from minor to significant is: sample tray temperature (ST20 and ST5) < mobile phase pH 9.5 (pH 9.5) ≈ column temperature 10^{o}C (CT10) < using different columns (used and Hused) << mobile phase pH 8.5 (pH 8.5) ≈ column temperature 40^{o}C (CT40). In general according to the

theoretical basis of reversed phase liquid chromatography, high temperature will increase the diffusion rates of solutes thus reducing partitioning into the stationary phase and thus reducing their retention times. However, this is not always the case in HILIC mode. [32] In the case of the two examples shown in Figure 4B, marked retention time changes in opposite directions could be observed for different metabolites under the conditions of CT40 (red marks) and pH 8.5 (blue marks). The detailed chromatographic mechanisms behind this observation will not be discussed here but will be explored in a subsequent study. Excluding these two extreme conditions, only 6 out of 400 individual LC-MS signals showed RSD in their retention times > ± 5% with the a maximum of ± 5.57% indicating good robustness of this HILIC method and most valuably only a small variation in retention times (4 out of 400 showed RSD > ± 5% with a maximum of ± 5.46%) was achieved on three different ZICpHILIC columns (ESI.5). Unlike column temperature and mobile phase pH, the reproducibility of chromatographic performance between different batch-numbered commercial HILIC columns is beyond the control of the researchers who perform LC-MS experiments. The current observations provide evidence that highly reproducible chromatographic performance can be achieved on the ZICpHILIC phase over multiple columns thus allowing alignment of chromatographic data run on columns columns with different batch numbers without applying retention time normalisation algorithms.

4 Conclusion

There is a lack of clear protocols for assessing the reliability and robustness of LC-MS methods for untargeted metabolomics. It is crucial that such methods are reliable and robust otherwise bioinformatics methods applied to the data generated will simply highlight unstable elements in the methodology due to variations in sample preparation,

chromatography and MS instrument response. In this study, human urine as matrix, we investigated the reliability and suitability of a HILIC-HRMS (ZIC-pHILIC-Exactive Orbitrap) method for untargeted metabolite profiling of human urine. From the study it is possible to summarise some key points with regard to achieving reliable data with this method. The HILIC-HRMS method required four QC injections to reach equilibrium conditions. Data variability introduced by sample preparation is larger than that produced during data acquisition by HILIC-HRMS for low-abundance urinary components. Therefore, multiple QC samples should be prepared from a single pooled sample and measured in parallel with real samples to evaluate the variations introduced in sample preparation [31],this is more even necessary for clinical samples derived from blood and tissue because more complicated sample preparation is involved. [30, 33] The subtle variation in datasets determined by the small fold changes of spiked isotope labelled amino acids was accurately observed by an OPLS model indicating the power of orthogonal multivariate treatments for observing subtle differences among samples. However, the analyst should be aware of potential matrix suppression effects on low abundance components. The pH of the mobile phase only has a strong effect if it falls below pH 9.0 and thus pH unadjusted ammonium carbonate which has a pH of *ca* 9.1 can be used. Column temperature should be controlled for this HILIC-HRMS method because dramatic changes in retention times were observed for some metabolites with elevated column temperature. It was demonstrated that highly reproducible chromatographic performance can be achieved on this polymer based zwitterionic stationary phase in HILIC mode between different columns thus potentially allowing data combination following column replacement in long term inter-batch analysis.

Acknowledgment

We acknowledge Scottish Life Sciences Alliance for support for mass spectrometry equipment.

References

[1] J.L. Spratlin, N.J. Serkova, S.G. Eckhardt, Clinical applications of metabolomics in oncology: a review, Clinical cancer research : an official journal of the American Association for Cancer Research, 15 (2009) 431-440.

[2] A. Roux, D. Lison, C. Junot, J.F. Heilier, Applications of liquid chromatography coupled to mass spectrometry-based metabolomics in clinical chemistry and toxicology: A review, Clinical biochemistry, 44 (2011) 119-135.

[3] S. Becker, L. Kortz, C. Helmschrodt, J. Thiery, U. Ceglarek, LC-MS-based metabolomics in the clinical laboratory, Journal of chromatography. B, Analytical technologies in the biomedical and life sciences, 883-884 (2012) 68-75.

[4] A. D'Alessandro, B. Giardina, F. Gevi, A.M. Timperio, L. Zolla, Clinical metabolomics: the next stage of clinical biochemistry, Blood transfusion = Trasfusione del sangue, 10 Suppl 2 (2012) s19-24.

[5] C. Baumgartner, A Clinical Metabolomics Strategy to Discover New Biomarkers in Complex Disease: An Overview, Biomedizinische Technik. Biomedical engineering, DOI 10.1515/bmt-2013-4305(2013).

[6] T. Zhang, D.G. Watson, A short review of applications of liquid chromatography mass spectrometry based metabolomics techniques to the analysis of human urine, The Analyst, 140 (2015) 2907-2915.

[7] H. Peretz, D.G. Watson, G. Blackburn, T. Zhang, A. Lagziel, M. Shtauber-Naamati, T. Morad, E. Keren-Tardai, V. Greenshpun, S. Usher, H. Shalev, D. Landau, D. Levartovsky, Urine metabolomics reveals novel physiologic functions of human aldehyde oxidase and provides biomarkers for typing xanthinuria, Metabolomics, 8 (2012) 951-959.

[8] M. Mraz, O. Hurba, J. Bartl, Z. Dolezel, A. Marinaki, L. Fairbanks, B. Stiburkova, Modern diagnostic approach to hereditary xanthinuria, Urolithiasis, 43 (2015) 61-67.

[9] D.K. Trivedi, R.K. Iles, Do not just do it, do it right: urinary metabolomics -establishing clinically relevant baselines, Biomed Chromatogr, 28 (2014) 1491-1501.

[10] M.M. Khamis, D.J. Adamko, A. El-Aneed, Mass spectrometric based approaches in urine metabolomics and biomarker discovery, Mass Spectrom Rev, DOI 10.1002/mas.21455(2015).

[11] A. Sreekumar, L.M. Poisson, T.M. Rajendiran, A.P. Khan, Q. Cao, J. Yu, B. Laxman, R. Mehra, R.J. Lonigro, Y. Li, M.K. Nyati, A. Ahsan, S. Kalyana-Sundaram, B. Han, X. Cao, J. Byun, G.S. Omenn, D. Ghosh, S. Pennathur, D.C. Alexander, A. Berger, J.R. Shuster, J.T. Wei, S. Varambally, C. Beecher, A.M. Chinnaiyan, Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression, Nature, 457 (2009) 910-914.

[12] N. Cernei, Z. Heger, J. Gumulec, O. Zitka, M. Masarik, P. Babula, T. Eckschlager, M. Stiborova, R. Kizek, V. Adam, Sarcosine as a potential prostate cancer biomarker--a review, International journal of molecular sciences, 14 (2013) 13893-13908.

[13] H.G. Gika, G.A. Theodoridis, R.S. Plumb, I.D. Wilson, Current practice of liquid chromatography-mass spectrometry in metabolomics and metabonomics, J Pharm Biomed Anal, 87 (2014) 12-25.

[14] H.G. Gika, I.D. Wilson, G.A. Theodoridis, LC-MS-based holistic metabolic profiling. Problems, limitations, advantages, and future perspectives, Journal of chromatography. B, Analytical technologies in the biomedical and life sciences, 966 (2014) 1-6.

[15] P.D. Rainville, G. Theodoridis, R.S. Plumb, I.D. Wilson, Advances in liquid chromatography coupled to mass spectrometry for metabolic phenotyping, Trac-Trend Anal Chem, 61 (2014) 181-191.

[16] T. Zhang, D.J. Creek, M.P. Barrett, G. Blackburn, D.G. Watson, Evaluation of Coupling Reversed Phase, Aqueous Normal Phase, and Hydrophilic Interaction Liquid Chromatography with Orbitrap Mass Spectrometry for Metabolomic Studies of Human Urine, Anal Chem, 84 (2012) 1994-2001.

[17] R. Zhang, D.G. Watson, L. Wang, G.D. Westrop, G.H. Coombs, T. Zhang, Evaluation of mobile phase characteristics on three zwitterionic columns in hydrophilic interaction liquid chromatography mode for liquid chromatography-high resolution mass spectrometry based untargeted metabolite profiling of Leishmania parasites, J Chromatogr A, 1362 (2014) 168-179.

[18] K. Contrepois, L. Jiang, M. Snyder, Optimized Analytical Procedures for the Untargeted Metabolomic Profiling of Human Urine and Plasma by Combining Hydrophilic Interaction and Reverse-Phase Liquid Chromatography - Mass Spectrometry, Molecular & cellular proteomics : MCP, DOI 10.1074/mcp.M114.046508(2015).

[19] S.U. Bajad, W. Lu, E.H. Kimball, J. Yuan, C. Peterson, J.D. Rabinowitz, Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry, Journal of chromatography. A, 1125 (2006) 76-88.

[20] D.P. Kloos, H. Lingeman, W.M. Niessen, A.M. Deelder, M. Giera, O.A. Mayboroda, Evaluation of different column chemistries for fast urinary metabolic profiling, Journal of chromatography. B, Analytical technologies in the biomedical and life sciences, 927 (2013) 90-96.

[21] D.L. Callahan, D. De Souza, A. Bacic, U. Roessner, Profiling of polar metabolites in biological extracts using diamond hydride-based aqueous normal phase chromatography, J Sep Sci, 32 (2009) 2273-2280.

[22] H.G. Gika, G.A. Theodoridis, J.E. Wingate, I.D. Wilson, Within-day reproducibility of an HPLC-MS-Based method for metabonomic analysis: Application to human urine, J Proteome Res, 6 (2007) 3291-3303.

[23] T.J. Demetrowitsch, B. Petersen, J.K. Keppler, A. Koch, S. Schreiber, M. Laudes, K. Schwarz, Validation of a two-step quality control approach for a large-scale human urine metabolomic study conducted in seven experimental batches with LC/QTOF-MS, Bioanalysis, 7 (2015) 103-112.

[24] P.A. Guy, I. Tavazzi, S.J. Bruce, Z. Ramadan, S. Kochhar, Global metabolic profiling analysis on human urine by UPLC-TOFMS: issues and method validation in nutritional metabolomics, Journal of chromatography. B, Analytical technologies in the biomedical and life sciences, 871 (2008) 253-260.

[25] E.J. Want, I.D. Wilson, H. Gika, G. Theodoridis, R.S. Plumb, J. Shockcor, E. Holmes, J.K. Nicholson, Global metabolic profiling procedures for urine using UPLC-MS, Nature protocols, 5 (2010) 1005-1018.

[26] D.Q. Tang, L. Zou, X.X. Yin, C.N. Ong, HILIC-MS for metabolomics: An attractive and complementary approach to RPLC-MS, Mass Spectrom Rev, DOI 10.1002/mas.21445(2014).

[27] E. Daskalaki, G. Blackburn, G. Kalna, T. Zhang, N. Anthony, D.G. Watson, A study of the effects of exercise on the urinary metabolome using normalisation to individual metabolic output, Metabolites, 5 (2015) 119-139.

[28] T. Pluskal, S. Castillo, A. Villar-Briones, M. Oresic, MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data, BMC bioinformatics, 11 (2010) 395.

[29] T. Zhang, D.G. Watson, L.J. Wang, M. Abbas, L. Murdoch, L. Bashford, I. Ahmad, N.Y. Lam, A.C.F. Ng, H.Y. Leung, Application of Holistic Liquid Chromatography-High Resolution Mass Spectrometry

Based Urinary Metabolomics for Prostate Cancer Detection and Biomarker Discovery, Plos One, 8 (2013).

[30] W.B. Dunn, D. Broadhurst, P. Begley, E. Zelena, S. Francis-McIntyre, N. Anderson, M. Brown, J.D. Knowles, A. Halsall, J.N. Haselden, A.W. Nicholls, I.D. Wilson, D.B. Kell, R. Goodacre, Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry, Nature protocols, 6 (2011) 1060-1083.

[31] W.B. Dunn, I.D. Wilson, A.W. Nicholls, D. Broadhurst, The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans, Bioanalysis, 4 (2012) 2249-2264.

[32] Z.G. Hao, B.M. Xiao, N.D. Weng, Impact of column temperature and mobile phase components on selectivity of hydrophilic interaction chromatography (HILIC), J Sep Sci, 31 (2008) 1449-1464.

[33] H. Gika, G. Theodoridis, Sample preparation prior to the LC-MS-based metabolomics/metabonomics of blood-derived samples, Bioanalysis, 3 (2011) 1647-1661.

Figure legends

Figure.1 Preparation of a series of urine samples for HILIC-HRMS method reliability and robustness evaluation. One of the QC-200 samples was used for QC injection (named as QCI), 7 times before running other samples and followed by every 10 samples and 5 times again after the whole sequence. DF=dilution factor, TV=total volume, AA=amino acid and ACN=acetonitrile.

Figure.2 HILIC-HRMS method technical variation evaluation using the samples prepared with varied urine volumes and dilution factors. The colour keys are shown at top-right in each figure and described in section 2.2 and Figure.1. A: PCA score plot of all the urine samples and the QC injections (QCI). B: PCA score plot of without the first four QCIs. C: PCA score plot of only the samples with urine dilution factor 5 including QCI5-17. D: column chart of the numbers of the LC-MS signals with RSD beyond 25% in peak area generated by QCIs, 200-QC samples and all the samples with urine dilution factor 5 including 50, 400 and 200-QC.

Figure.3 HILIC-HRMS method reliability evaluation using the samples with varied concentrations of ^{13}C and ^{15}N isotope labelled amino acids. The colour keys are shown at top-right in each figure and described in section 2.1 and Figure.1. A: PCA score plot of the urine samples with and without spiked isotope labelled amino acids. B: OPLS score plot of the same samples in A. C: PCA score plot of water samples with spiked isotope labelled amino acids. D: OPLS loading plot of urinary metabolites vs. dilution factors of spiked isotope labelled amino acids (X variables: metabolites and Y variables: dilution factors). The sixteen identified isotope labelled amino acids (in both ESI positive and negative modes) are marked in solid red and the other three metabolites mentioned in the text are marked in solid green and are clearly orthogonal to the specified Y variables since they are also present in the urinary matrix.

Figure.4 HILIC-HRMS method robustness test: LC parameters (see Table.1). A: PCA score plot of the replicate injections with varied LC parameters. The colour keys are shown at top-right in the figure. B: Two examples for the metabolites showing dramatic but opposite changes in retention time under the conditions of CT40 (in red) and pH8.5 (in blue).

***Highlights (for review)**

Highlights

1. The technical variations introduced by sample preparation was found more significant than the current HILIC-HRMS platform after multivariate (PCA) and univariate (RSD) analysis of the urine samples (injections) prepared with varied urine volumes and dilution factors.
2. The suitability of the current HILIC-HRMS method to clinical metabolomics was demonstrated by supervised multivariate analysis (OPLS model) of the urine samples spiked with isotope labelled amino acids at varied concentrations.
3. Several key LC parameters were investigated in respect of LC stability for inter-batch analysis and high reproducibility of chromatographic performance was observed with this polymer based zwitterionic stationary phase in HILIC mode.

Figure 1

Click here to download high resolution image

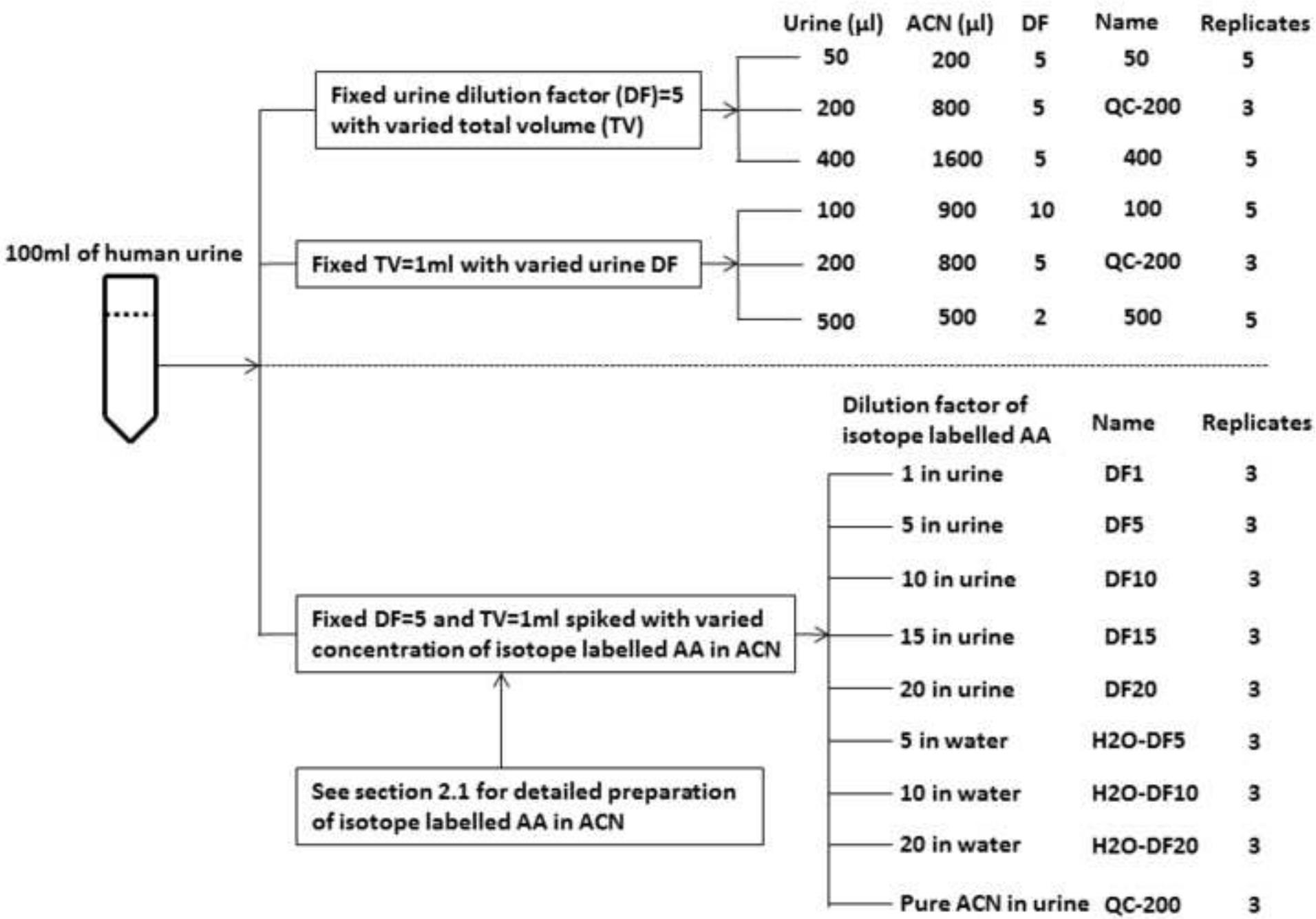


Figure 2

Click here to download high resolution image

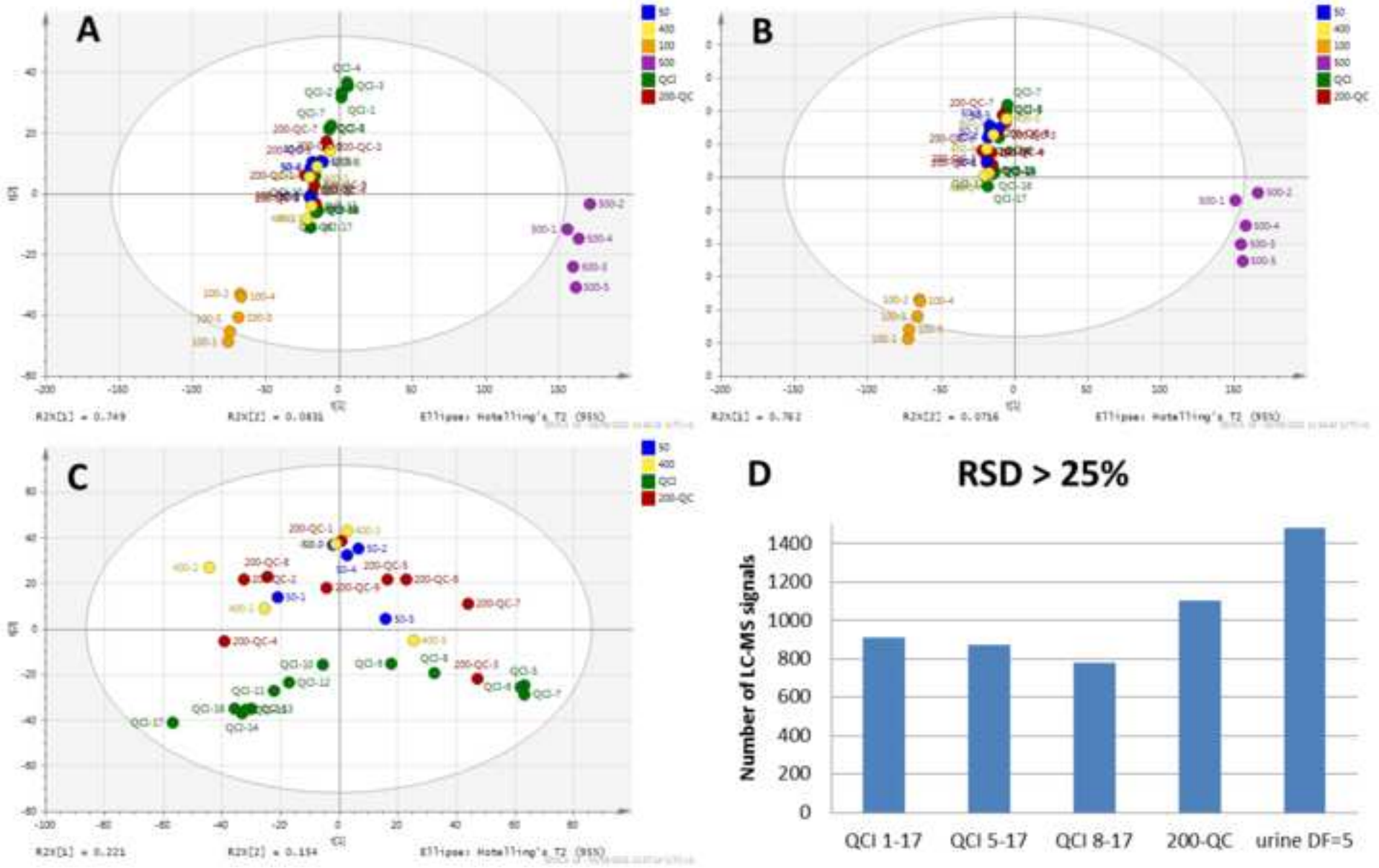


Figure 3

Click here to download high resolution image

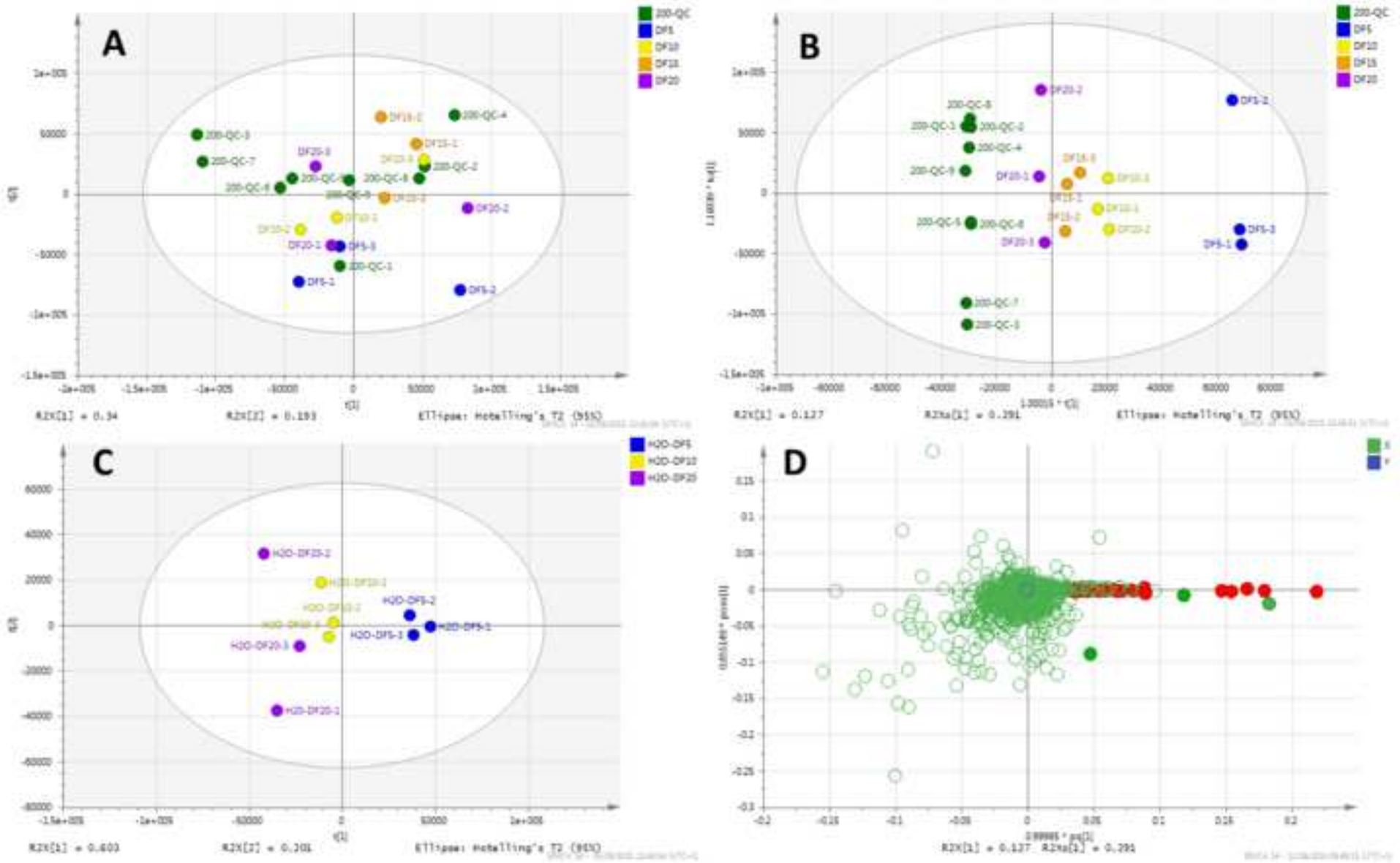
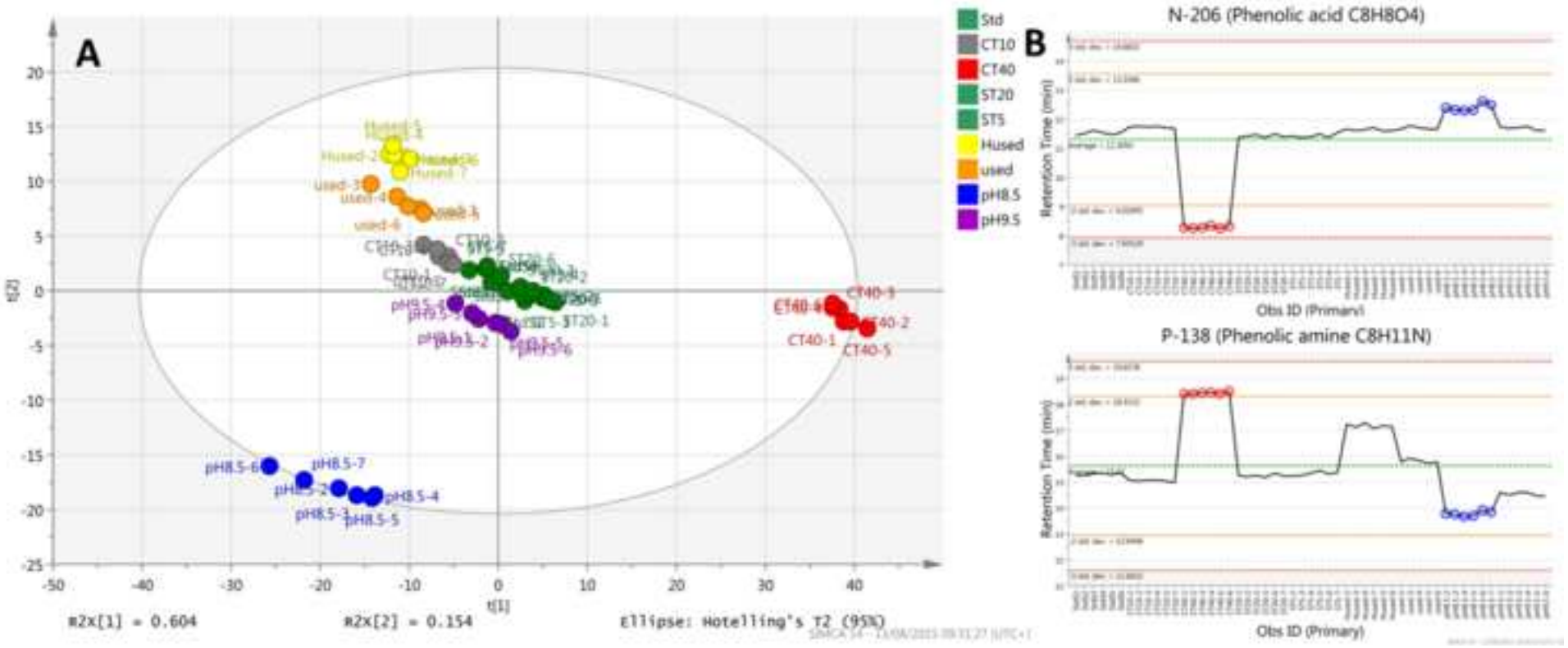


Figure 4

Click here to download high resolution image



Electronic Supplementary Information 1

Click here to download Supplementary material for on-line publication only: ESI.1.pdf

Electronic Supplementary Information 2

Click here to download Supplementary material for on-line publication only: ESI.2.xlsx

Electronic Supplementary Information 3

Click here to download Supplementary material for on-line publication only: ESI.3.pdf

Electronic Supplementary Information 4

Click here to download Supplementary material for on-line publication only: ESI.4.pdf

Electronic Supplementary Information 5

Click here to download Supplementary material for on-line publication only: ESI.5.xlsx